

A RELATIVE VALUE ITERATION ALGORITHM FOR NON-DEGENERATE CONTROLLED DIFFUSIONS

ARI ARAPOSTATHIS* AND VIVEK S. BORKAR†

Abstract. The ergodic control problem for a non-degenerate controlled diffusion controlled through its drift is considered under a uniform stability condition that ensures the well-posedness of the associated Hamilton–Jacobi–Bellman (HJB) equation. A nonlinear parabolic evolution equation is then proposed as a continuous time continuous state space analog of White’s ‘*relative value iteration*’ algorithm for solving the ergodic dynamic programming equation for the finite state finite action case. Its convergence to the solution of the HJB equation is established using the theory of monotone dynamical systems and also, alternatively, by using the theory of reverse martingales.

Key words. controlled diffusions; ergodic control; Hamilton–Jacobi–Bellman equation; relative value iteration; monotone dynamical systems; reverse martingales

AMS subject classifications. Primary, 93E15, 93E20; Secondary, 60J25, 60J60, 90C40

1. Introduction. Consider a controlled Markov chain on a finite state space $\mathcal{S} = \{1, \dots, N\}$ with transition probabilities $p_{ij}(u)$, $i, j \in \mathcal{S}$, which depend continuously on a control parameter u that lives in a compact ‘action’ space \mathbb{U} , such that when in state i the control u is chosen from a compact subset $\mathbb{U}_i \subset \mathbb{U}$. Assuming irreducibility for the stochastic matrix $P^v \triangleq [p_{i,j}(v_i)]_{i,j \in \mathcal{S}}$ with $v = (v_1, \dots, v_N) \in (\mathbb{U}_1 \times \dots \times \mathbb{U}_N)$, consider the control problem of minimizing the average (or ergodic) cost

$$\limsup_{n \uparrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[r(X_k, U_k)]$$

for a prescribed $r : \mathcal{S} \times \mathbb{U} \rightarrow \mathbb{R}$ and control sequence $\{U_k\}$ such that $U_k \in \mathbb{U}_{X_k}$ and

$$\mathbb{P}(X_{n+1} = j \mid X_m, U_m, m \leq n) = p_{X_n j}(U_n), \quad n \geq 0.$$

The dynamic programming equation for this problem is the well known controlled Poisson equation:

$$V(i) = \min_{u \in \mathbb{U}_i} \left[r(i, u) - \beta + \sum_{j \in \mathcal{S}} p_{ij}(u) V(j) \right], \quad i \in \mathcal{S}.$$

This is an equation in unknowns (V, β) , with $V = (V(1), \dots, V(N)) \in \mathbb{R}^N$ the so called value function. Under the irreducibility hypothesis above, V is uniquely specified modulo an additive constant and β is uniquely specified as the optimal ergodic cost. See [DY79, Put94] for details.

*Department of Electrical and Computer Engineering, The University of Texas at Austin, 1 University Station, Austin, TX 78712 (ari@mail.utexas.edu). This author’s work was supported in part by the Office of Naval Research through the Electric Ship Research and Development Consortium.

†Department of Electrical Engineering, Indian Institute of Technology, Powai, Mumbai 400076, India. (borkar.vs@gmail.com). This author’s work was supported in part by the J. C. Bose Fellowship from the Government of India.

By analogy with the value iteration algorithm for the discounted cost problem, one may consider the value iteration algorithm

$$V^{n+1}(i) = \min_{u \in \mathbb{U}_i} \left[r(i, u) - \beta + \sum_{j \in \mathcal{S}} p_{ij}(u) V^n(j) \right], \quad i \in \mathcal{S}, \quad (1.1)$$

beginning with an initial guess $V^0(\cdot)$. The difficulty here is that β is unknown as well. On the other hand, if we drop β from (1.1), there is no convergence—the map $V^n \mapsto V^{n+1} \triangleq F(V^n)$ that is being iterated lacks the contractivity property of its discounted cost counterpart. Thus clearly some renormalization is required. The earliest example of such a *relative value iteration* algorithm for finite state Markov chains is perhaps that of White [Whi63], which is governed by

$$h_{k+1}(i) = \min_{u \in \mathbb{U}_i} \left[r(i, u) + \sum_{j=1}^n p_{ij}(u) h_k(j) \right] - \lambda_{k+1} \quad (1.2a)$$

$$\lambda_{k+1} = \min_{u \in \mathbb{U}_n} \left[r(n, u) + \sum_{j=1}^n p_{nj}(u) h_k(j) \right]. \quad (1.2b)$$

For a discussion of other possible choices for updating (1.2b) see [ABB01].

Bertsekas introduced in [Ber98] a variation of this method that takes the form

$$h_{k+1}(i) = \min_{u \in \mathbb{U}_i} \left[r(i, u) + \sum_{j=1}^{n-1} p_{ij}(u) h_k(j) \right] - \lambda_k$$

$$\lambda_{k+1} = \lambda_k + \gamma_k h_{k+1}(n).$$

Here $\{\gamma_k\}$ is a sequence of positive stepsizes. This has led to the learning algorithms analyzed in [ABB01]. Recently Shlakhter et. al. [SLKJ10] have studied ways of accelerating the convergence of the above value iteration algorithms.

Studies of convergence of relative value iteration schemes for more general Markov processes are non-existent. The only related work that comes to mind is convergence of the value iteration in (1.1) for denumerable controlled Markov chains [AF99].

Our aim in this paper is to propose a relative value iteration scheme in continuous time and space for a class of controlled diffusion processes and prove its convergence. While we prefer to think of this scheme as a continuous time and space relative value iteration, it can also be viewed as a ‘stabilization of a nonlinear parabolic PDE problem in the sense of Has’minskiĭ (see [Has60]). We follow two different approaches for the proof of convergence, based on resp. the theory of monotone dynamical systems and the theory of reverse martingales. These should be of independent interest.

The paper is organized as follows. The next section describes the ergodic control problem for diffusions and the associated Hamilton–Jacobi–Bellman equation, leading to the proposed relative value iteration scheme. Section 3 provides a motivating illustration from the discrete state counterpart, introduces some notation, and recalls some key results from parabolic PDEs and monotone dynamical systems for later use. Section 4 gives the two convergence proofs alluded in the Abstract, while Section 5 concludes with some pointers to future work.

2. Problem statement.

2.1. The model. We are concerned with controlled diffusion processes $X = \{X_t, t \geq 0\}$ taking values in the d -dimensional Euclidean space \mathbb{R}^d , and governed by the Itô stochastic differential equation

$$dX_t = b(X_t, U_t) dt + \sigma(X_t) dW_t. \quad (2.1)$$

All random processes in (2.1) live in a complete probability space $(\Omega, \mathfrak{F}, \mathbb{P})$. The process W is a d -dimensional standard Wiener process independent of the initial condition X_0 . The control process U takes values in a compact, metrizable set \mathbb{U} , and $U_t(\omega)$ is jointly measurable in $(t, \omega) \in [0, \infty) \times \Omega$. Moreover, it is *non-anticipative*: for $s < t$, $W_t - W_s$ is independent of

$$\mathfrak{F}_s \triangleq \text{the completion of } \sigma\{X_0, U_r, W_r, r \leq s\} \text{ relative to } (\mathfrak{F}, \mathbb{P}).$$

Such a process U is called an *admissible control*, and we let \mathfrak{U} denote the set of all admissible controls.

We impose the following standard assumptions on the drift b and the diffusion matrix σ to guarantee existence and uniqueness of solutions to (2.1).

(A1) Local Lipschitz continuity: The functions

$$b = [b^1, \dots, b^d]^\top : \mathbb{R}^d \times \mathbb{U} \mapsto \mathbb{R}^d \quad \text{and} \quad \sigma = [\sigma^{ij}] : \mathbb{R}^d \mapsto \mathbb{R}^{d \times d}$$

are locally Lipschitz in x with a Lipschitz constant κ_R depending on $R > 0$. In other words, if B_R denotes the open ball of radius R centered at the origin in \mathbb{R}^d , then for all $x, y \in B_R$ and $u \in \mathbb{U}$,

$$|b(x, u) - b(y, u)| + \|\sigma(x) - \sigma(y)\| \leq \kappa_R |x - y|,$$

where $\|\sigma\|^2 \triangleq \text{trace}(\sigma\sigma^\top)$.

(A2) Affine growth condition: b and σ satisfy a global growth condition of the form

$$|b(x, u)|^2 + \|\sigma(x)\|^2 \leq \kappa_1 (1 + |x|^2) \quad \forall (x, u) \in \mathbb{R}^d \times \mathbb{U}.$$

(A3) Local non-degeneracy: Let $a \triangleq \frac{1}{2} \sigma \sigma^\top$. For each $R > 0$, we have

$$\sum_{i,j=1}^d a^{ij}(x) \xi_i \xi_j \geq \kappa_R^{-1} |\xi|^2 \quad \forall x \in B_R,$$

for all $\xi = (\xi_1, \dots, \xi_d) \in \mathbb{R}^d$.

We also assume that b is continuous in (x, u) .

In integral form, (2.1) is written as

$$X_t = X_0 + \int_0^t b(X_s, U_s) ds + \int_0^t \sigma(X_s) dW_s. \quad (2.2)$$

The second term on the right hand side of (2.2) is an Itô stochastic integral. We say that a process $X = \{X_t(\omega)\}$ is a solution of (2.1), if it is \mathfrak{F}_t -adapted, continuous in t , defined for all $\omega \in \Omega$ and $t \in [0, \infty)$, and satisfies (2.2) for all $t \in [0, \infty)$ at once a.s.

With $u \in \mathbb{U}$ treated as a parameter, we define the family of operators $L^u : \mathcal{C}^2(\mathbb{R}^d) \mapsto \mathcal{C}(\mathbb{R}^d)$ by

$$L^u f(x) = \sum_{i,j} a^{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_i b^i(x, u) \frac{\partial f}{\partial x_i}(x), \quad u \in \mathbb{U}. \quad (2.3)$$

We refer to L^u as the *controlled extended generator* of the diffusion.

Of fundamental importance in the study of functionals of X is Itô's formula. For $f \in \mathcal{C}^2(\mathbb{R}^d)$ and with L^u as defined in (2.3),

$$f(X_t) = f(X_0) + \int_0^t L^{U_s} f(X_s) ds + M_t, \quad \text{a.s.}, \quad (2.4)$$

where

$$M_t \triangleq \int_0^t \langle \nabla f(X_s), \sigma(X_s) dW_s \rangle$$

is a local martingale. Krylov's extension of the Itô formula [Kry80, p. 122] extends (2.4) to functions f in the local Sobolev space $\mathcal{W}_{\text{loc}}^{2,p}(\mathbb{R}^d)$.

Recall that a control is called *Markov* if $U_t = v(t, X_t)$ for a measurable map $v : \mathbb{R} \times \mathbb{R}^d \mapsto \mathbb{U}$, and it is called *stationary Markov* if v does not depend on t , i.e., $v : \mathbb{R}^d \mapsto \mathbb{U}$. Correspondingly, the equation

$$X_t = x_0 + \int_0^t b(X_s, v(s, X_s)) ds + \int_0^t \sigma(X_s) dW_s \quad (2.5)$$

is said to have a *strong solution* if given a Wiener process (W_t, \mathfrak{F}_t) on a complete probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, there exists a process X on $(\Omega, \mathfrak{F}, \mathbb{P})$, with $X_0 = x_0 \in \mathbb{R}^d$, which is continuous, \mathfrak{F}_t -adapted, and satisfies (2.5) for all t at once, a.s. A strong solution is called *unique*, if any two such solutions X and X' agree \mathbb{P} -a.s., when viewed as elements of $\mathcal{C}([0, \infty), \mathbb{R}^d)$. It is well known that under Assumptions (A1)–(A3), for any Markov control v , (2.5) has a unique strong solution [GK96].

Let \mathfrak{U}_{SM} denote the set of stationary Markov controls. Under $v \in \mathfrak{U}_{\text{SM}}$, the process X is strong Markov, and we denote its transition function by $P_v^t(x, \cdot)$. It also follows from the work of [BKR01, Sta99] that under $v \in \mathfrak{U}_{\text{SM}}$, the transition probabilities of X have densities which are locally Hölder continuous. Thus L^v defined by

$$L^v f(x) = \sum_{i,j} a^{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_i b^i(x, v(x)) \frac{\partial f}{\partial x_i}(x), \quad v \in \mathfrak{U}_{\text{SM}},$$

for $f \in \mathcal{C}^2(\mathbb{R}^d)$, is the generator of a strongly-continuous semigroup on $\mathcal{C}_b(\mathbb{R}^d)$, which is strong Feller. We let \mathbb{P}_x^v denote the probability measure and \mathbb{E}_x^v the expectation operator on the canonical space of the process under the control $v \in \mathfrak{U}_{\text{SM}}$, conditioned on the process X starting from $x \in \mathbb{R}^d$ at $t = 0$.

2.2. The ergodic control problem. Let $r : \mathbb{R}^d \times \mathbb{U} \rightarrow \mathbb{R}$ be a continuous function bounded from below, referred to as the *running cost*. As is well known, the ergodic control problem, in its *almost sure* (or *pathwise*) formulation, seeks to a.s. minimize over all admissible $U \in \mathfrak{U}$

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t r(X_s, U_s) ds. \quad (2.6)$$

A weaker, *average* formulation seeks to minimize

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t \mathbb{E}^U [r(X_s, U_s)] \, ds. \quad (2.7)$$

We let β be defined as

$$\beta \triangleq \inf_{U \in \mathfrak{U}} \limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t \mathbb{E}^U [r(X_s, U_s)] \, ds, \quad (2.8)$$

i.e., the infimum of (2.7) over all admissible controls.

We assume that the running cost function $r: \mathbb{R}^d \times \mathbb{U} \rightarrow \mathbb{R}_+$ is continuous and locally Lipschitz in its first argument uniformly in $u \in \mathbb{U}$. Without loss of generality we let κ_R be a Lipschitz constant of r over B_R , i.e., More specifically, for some function $K_c: \mathbb{R}_+ \rightarrow \mathbb{R}_+$,

$$|r(x, u) - r(y, u)| \leq \kappa_R |x - y| \quad \forall x, y \in B_R, \quad \forall u \in \mathbb{U},$$

and all $R > 0$.

We work under the following stability assumption:

Assumption 2.1. *There exists a nonnegative, inf-compact $\mathcal{V}: \mathbb{R}^d \rightarrow \mathbb{R}$ and positive constants c_0 , c_1 and c_2 satisfying*

$$L^u \mathcal{V}(x) \leq c_0 - c_1 \mathcal{V}(x) \quad \forall u \in \mathbb{U} \quad (2.9a)$$

$$\sup_{u \in \mathbb{U}} r(x, u) \leq c_2 \mathcal{V}(x) \quad (2.9b)$$

for all $x \in \mathbb{R}^d$. Without loss of generality we assume $\mathcal{V} \geq 1$.

It is well known (see [ABG11, GS72]) that (2.9a) implies that

$$\mathbb{E}_x^U [\mathcal{V}(X_t)] \leq \frac{c_0}{c_1} + \mathcal{V}(x) e^{-c_1 t} \quad \forall x \in \mathbb{R}^d, \quad \forall U \in \mathfrak{U}. \quad (2.10)$$

Recall that control $v \in \mathfrak{U}_{\text{SM}}$ is called *stable* if the associated diffusion is positive recurrent. We denote the set of such controls by $\mathfrak{U}_{\text{SSM}}$. Also we let μ_v denote the unique invariant probability measure on \mathbb{R}^d for the diffusion under the control $v \in \mathfrak{U}_{\text{SSM}}$. It follows by (2.10) that, under Assumption 2.1, all stationary Markov controls are stable and that

$$\int_{\mathbb{R}^d} \mathcal{V}(x) \mu_v(dx) \leq \frac{c_0}{c_1}.$$

Let $\mathcal{C}_{\mathcal{V}}(\mathbb{R}^d)$ denote the Banach space of functions in $\mathcal{C}(\mathbb{R}^d)$ with norm $\|f\|_{\mathcal{V}} \triangleq \sup_{x \in \mathbb{R}^d} \left| \frac{f(x)}{\mathcal{V}(x)} \right|$. Recall that a skeleton of a continuous-time Markov process is a discrete-time Markov process with transition probability $\tilde{P} = \int_0^\infty \alpha(dt) P^t$, where α is a probability measure on $(0, \infty)$. Since the diffusion is non-degenerate, any skeleton of the process is ϕ -irreducible, with an irreducibility measure absolutely continuous with respect to the Lebesgue measure. It is also straightforward to show that compact subsets of \mathbb{R}^d are petite. It then follows that for any $v \in \mathfrak{U}_{\text{SSM}}$ the controlled process under v is \mathcal{V} -geometrically ergodic (see [DMT95, FR05]), or in other words there exist constants C_0 and $\gamma > 0$ such that if $h \in \mathcal{C}_{\mathcal{V}}(\mathbb{R}^d)$ then

$$\left| P_v^t h(x) - \int_{\mathbb{R}^d} h(x) \mu_v(dx) \right| \leq C_0 e^{-\gamma t} \|h\|_{\mathcal{V}} \mathcal{V}(x), \quad t \geq 0, \quad x \in \mathbb{R}^d.$$

Concerning the ergodic control problem the following result is standard [ABG11].

THEOREM 2.2. *Under Assumption 2.1 there exists a unique solution $V^* \in \mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$, satisfying $V^*(0) = 0$, of*

$$0 = \min_{u \in \mathbb{U}} [L^u V^*(x) + r(x, u)] - \beta. \quad (2.11)$$

A control $v^* \in \mathfrak{U}_{\text{SM}}$ is optimal with respect to the criteria (2.6) and (2.7) if and only if it satisfies

$$\min_{u \in \mathbb{U}} \left[\sum_{i=1}^d b^i(x, u) \frac{\partial V}{\partial x_i}(x) + r(x, u) \right] = \sum_{i=1}^d b^i(x, v^*(x)) \frac{\partial V}{\partial x_i}(x) + r(x, v^*(x)) \quad (2.12)$$

a.e. in \mathbb{R}^d .

For the rest of the paper $v^* \in \mathfrak{U}_{\text{SSM}}$ denotes some fixed control satisfying (2.12).

2.3. The relative value iteration. We study the following relative value iteration (RVI) scheme:

$$\frac{\partial V}{\partial t}(t, x) = \min_{u \in \mathbb{U}} [L^u V(t, x) + r(x, u)] - V(t, 0), \quad V(0, x) = V_0(x), \quad (2.13)$$

with the boundary condition $V_0 \in \mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$.

The main theorem of the paper is as follows.

THEOREM 2.3. *For each $V_0 \in \mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$, the solution $V(t, x)$ of (2.13) converges to $V^*(x) + \beta$ as $t \rightarrow \infty$.*

The proof of convergence of (2.13) is facilitated by the study of the *value iteration* (VI) equation

$$\frac{\partial \bar{V}}{\partial t}(t, x) = \min_{u \in \mathbb{U}} [L^u \bar{V}(t, x) + r(x, u)] - \beta, \quad \bar{V}(0, x) = V_0(x). \quad (2.14)$$

Here $V_0 \in \mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$ as in (2.13). Also β is as in (2.8), so it is assumed known.

As shown in Lemma 4.2 in Section 4, $\bar{V}(t, \cdot)$ is bounded in $\mathcal{C}_V(\mathbb{R}^d)$ uniformly in $t \geq 0$. By (2.14) we have

$$\bar{V}(t, x) = \inf_{U \in \mathfrak{U}} \left(\int_0^t \mathbb{E}_x^U [r(X_s, U_s) - \beta] ds + \mathbb{E}_x^U [V_0(X_t)] \right). \quad (2.15)$$

Also, as we show in Lemma 4.4,

$$V(t, x) = \bar{V}(t, x) - e^{-t} \int_0^t e^s \bar{V}(s, 0) ds + \beta(1 - e^{-t}) \quad \forall x \in \mathbb{R}^d, t \geq 0.$$

It follows that $V(t, \cdot)$ is also bounded in $\mathcal{C}_V(\mathbb{R}^d)$ uniformly in $t \geq 0$. Additionally, convergence of $\bar{V}(t, \cdot)$ as $t \rightarrow \infty$ implies the analogous convergence of $V(t, \cdot)$. In Section 4 we provide two separate proofs of convergence of $\bar{V}(t, \cdot)$ as $t \rightarrow \infty$ to a solution of (2.11). The first employs results from the theory of monotone dynamical systems, while the second utilizes a reverse martingale convergence theorem.

Remark 2.4. Note that by (2.15) convergence of $\bar{V}(t, \cdot)$ as $t \rightarrow \infty$ to a solution of (2.11) implies that $F(t, \cdot)$ defined by

$$F(t, x) \triangleq \inf_{U \in \mathfrak{U}} \int_0^t \mathbb{E}_x^U [r(X_s, U_s) - \beta] ds$$

also converges to a solution of the HJB equation in (2.11).

Note also that the (VI) provides a sharp bound for the performance of an optimal ergodic control v^* over a finite horizon. Indeed, by (2.11), we have

$$V^*(x) = \mathbb{E}_x^{v^*} \left[\int_0^t (r(X_s, U_s) - \beta) \, ds \right] + \mathbb{E}_x^{v^*} [V^*(X_t)] .$$

Therefore, by (2.14) with boundary condition $V_0 \equiv 0$, we obtain

$$\int_0^t \mathbb{E}_x^{v^*} [r(X_t, U_t)] - \inf_{U \in \mathfrak{U}} \int_0^t \mathbb{E}_x^U [r(X_t, U_t)] = V^*(x) - \mathbb{E}_x^{v^*} [V^*(X_t)] - \bar{V}(t, x), \quad (2.16)$$

and the infimum is realized by any measurable selector from the minimizer of the (VI). Since the right hand side of (2.16) is bounded in $\mathcal{C}_V(\mathbb{R}^d)$ uniformly in $t \in [0, \infty)$, it follows that, under Assumption 2.1, a stationary Markov average-cost optimal control v^* satisfies

$$\int_0^T \mathbb{E}_x^{v^*} [r(X_t, U_t)] \, dt \leq K_0 \mathcal{V}(x) + \inf_{U \in \mathfrak{U}} \int_0^T \mathbb{E}_x^U [r(X_t, U_t)] \, dt \quad \forall T \geq 0 .$$

This provides a sharp bound for bias and overtaking optimality over the class of all Markov controls (compare with the results in [JFHL09] which are restricted to the class of optimal stationary Markov controls).

3. Preliminaries.

3.1. A Result from Monotone Dynamical Systems. Let \mathcal{H} be a subset of a metric space \mathcal{Y} of real valued functions defined on a set \mathcal{X} . Suppose also that \mathcal{H} is a subset of a Banach space \mathcal{G} with a positive cone \mathcal{G}_+ which has a nonempty interior. Let \preceq be the natural partial order on \mathcal{H} relative to the positive cone of \mathcal{G}_+ . In other words, for $h, h' \in \mathcal{H}$ we write $h \preceq h'$ if $h' - h \in \mathcal{G}_+$ for all $x \in \mathcal{X}$. We also introduce the relation \ll and write $h \ll h'$ if $h' - h \in \text{int}(\mathcal{G}_+)$, where ‘int’ denotes the interior.

Let $\Phi : \mathcal{H} \times \mathbb{R}_+ \rightarrow \mathcal{H}$ be a semiflow on \mathcal{H} . In other words, Φ satisfies

- (i) $\Phi_0(h) = h$ for all $h \in \mathcal{H}$;
- (ii) $\Phi_t \circ \Phi_s = \Phi_{t+s}$ for all $t, s \in \mathbb{R}_+$.

As well known, if $h \in \mathcal{H}$, then its *orbit* $O(h)$ is defined by $O(h) \triangleq \{\Phi_t(h) : t \geq 0\}$. Also the ω -*limit set* of $h \in \mathcal{H}$ is denoted by $\omega(h)$ and defined as $\omega(h) \triangleq \bigcap_{t>0} \overline{\bigcup_{s \geq t} \Phi_t(h)}$, where the closure is in \mathcal{Y} . The semiflow is called *monotone* (*strongly monotone*) if $h \preceq h'$ ($h \ll h'$) implies that $\Phi_t(h) \preceq \Phi_t(h')$ ($\Phi_t(h) \ll \Phi_t(h')$) for all $t > 0$. It is called *eventually strongly monotone* if it is monotone and whenever $h \prec h'$ there exists some $t_0 \in \mathbb{R}_+$ such that $\Phi_{t_0}(h) \ll \Phi_{t_0}(h')$. As shown in [Smi95, Proposition 1.1], if Φ is eventually strongly monotone then it is *strongly order preserving* (SOP), and this means that whenever $h \prec h'$ there exist open neighborhoods U and U' of h and h' , respectively, and $t_0 > 0$ such that $\Phi_t(U) \preceq \Phi_t(U')$ for all $t > t_0$.

Let

$$\mathcal{E} \triangleq \{h \in \mathcal{H} : \Phi_t(h) = h, \, \forall t \geq 0\} .$$

In other words, \mathcal{E} is the set of equilibria of the semiflow. A point $h \in \mathcal{H}$ is called *quasiconvergent* if $\omega(h) \subset \mathcal{E}$, and *convergent* if $\omega(h)$ is a singleton. Let \mathfrak{Q} and \mathfrak{C} denote the sets of quasiconvergent and convergent points, respectively.

We quote the following theorem [Smi95, Theorem 4.3 and Remark 4.2] which shows that quasi-convergence is generic. We need the following notation: We write $h_n \uparrow\uparrow h$ ($h_n \downarrow\downarrow h$) if $h_n \prec h_{n+1}$ ($h_n \succ h_{n+1}$) and $\lim_n h_n \rightarrow h$ in \mathcal{H} .

THEOREM 3.1. *Let Φ_t be a strongly preserving semiflow on $\mathcal{H} \subset \mathcal{G}$. Suppose that*

(i) *For any $h \in \mathcal{H}$ there exists a sequence $\{h_n\} \subset \mathcal{H}$ such that $h_n \uparrow\uparrow h$ or $h_n \downarrow\downarrow h$.*

(ii) *For each $h \in \mathcal{H}$ the closure of $O(h)$ is a compact subset of \mathcal{H} .*

(iii) *If $\{h_n\} \subset \mathcal{H}$ is such that $h_n \uparrow\uparrow h$ or $h_n \downarrow\downarrow h$, then $\{\cup_{n \in \mathbb{N}} \omega(h_n)\}$ has compact closure in \mathcal{Y} which is contained in \mathcal{H} .*

Then $\mathcal{H} = \text{int}(\mathfrak{Q}) \cup \overline{\text{int}(\mathfrak{C})}$. Moreover, if \mathcal{E} is totally ordered with respect to \preceq , then $\mathfrak{Q} = \mathfrak{C}$ which implies that $\mathcal{H} = \overline{\text{int}(\mathfrak{C})}$.

3.2. The case of continuous time controlled Markov chains. To illustrate our approach, we consider here the simple case of a controlled Markov chain with state space \mathcal{S} in continuous time, with ‘rate matrix’ $Q^u = [q_{ij}(u)]$, $i, j \in \mathcal{S}$, depending continuously on a parameter u that lives in a compact action space \mathbb{U} . The matrix Q satisfies $q_{ij} \geq 0$ for all $i \neq j$ and $\sum_{j \in \mathcal{S}} q_{ij} = 0$. Suppose first that the state space is finite, i.e., $\mathcal{S} = \{1, \dots, N\}$. To guarantee irreducibility we assume that there exists an irreducible rate matrix $\bar{Q} = [\bar{q}_{ij}]$ and a constant $\delta > 0$ such that $q_{ij}(u) \geq \delta \bar{q}_{ij}$ for all $i \neq j$ and $u \in \mathbb{U}$. Let $r : \mathcal{S} \times \mathbb{U} \rightarrow \mathbb{R}$ be a running cost. The solution of the ergodic control problem has the following characterization: There exists a unique pair (V^*, β) with β a constant and $V^* : \mathcal{S} \rightarrow \mathbb{R}$, satisfying $V^*(N) = 0$, which solve with $V = V^*$ the equation

$$\min_{u \in \mathbb{U}} \left[\sum_{j \in \mathcal{S}} q_{ij}(u) V(j) + r(i, u) \right] = \beta \quad \forall i \in \mathcal{S}. \quad (3.1)$$

Moreover a stationary Markov control $v = (v_1, \dots, v_N)$ is average-cost optimal if and only if it is a selector from the minimizer in (3.1). Expressing r in vector form as $r(u) = (r(1, u), \dots, r(N, u))^T$, the relative value iteration algorithm takes the form of the following differential equation in \mathbb{R}^N :

$$\frac{dh}{dt} = \min_{u \in \mathbb{U}} [Q(u)h + r(u)] - \mathbf{1}h_N(t), \quad h(0) = g \in \mathbb{R}^N. \quad (3.2)$$

where $\mathbf{1}$ indicates the vector whose components are all equal to 1. Showing existence of solutions to (3.2) is straightforward. One can follow for example the method used in the proof of Lemma 4.1 which appears in Section 4. The corresponding value iteration equation is

$$\frac{d\bar{h}}{dt} = \min_{u \in \mathbb{U}} [Q(u)\bar{h} + r(u)] - \mathbf{1}\beta, \quad h(0) = g \in \mathbb{R}^N. \quad (3.3)$$

We apply Theorem 3.1 to (3.3). Here \mathcal{H} and \mathcal{G} are isomorphic to \mathbb{R}^N under the Euclidean norm topology. Hence the partial ordering is $\bar{h} \preceq \bar{h}' \iff \bar{h}_i \leq \bar{h}'_i$ for all $i \in \mathcal{S}$. The fact that (3.3) is strongly order preserving follows from the irreducibility of the chain. Hypothesis (i) of Theorem 3.1 is obviously satisfied in $\mathcal{H} \sim \mathbb{R}^N$. Since the solution of (3.3) is uniformly bounded for any initial condition with the bound depending continuously on the initial condition g , it follows that hypotheses (ii) and (iii) of Theorem 3.1 are satisfied. The equilibrium set \mathcal{E} of (3.3) is the set of $V \in \mathbb{R}^N$ which solve (3.1). Hence $\mathcal{E} = \{V^* + c : c \in \mathbb{R}\}$, which is a totally ordered set. It then follows from Theorem 3.1 that $\mathcal{H} = \overline{\text{int}(\mathfrak{C})}$. It is also straightforward to show from (3.3) that the solutions are

continuous with respect to the initial condition, uniformly in $t \in [0, \infty)$, i.e., that if g^n is a sequence converging $g \in \mathbb{R}^N$ as $n \rightarrow \infty$, then

$$\sup_{t \geq 0} |\Phi_t(g^n) - \Phi_t(g)| \xrightarrow{n \rightarrow \infty} 0.$$

As a result, \mathfrak{C} is closed and hence every initial condition is convergent point. By (3.2)–(3.3) and following the argument at the end of Section 4.1 for the proof of Theorem 2.3, it follows that $h(t)$ converges to $V^* + \beta$. Convergence of the relative value iteration for countable state space Markov chains in continuous time follows along the same lines, provided a Lyapunov hypothesis analogous to (2.9a) is imposed, as well as appropriate assumptions to guarantee the regularity of the process. We don't delve into these details, since the focus in this paper is continuous state space models.

3.3. Notation and Background. The term *domain* in \mathbb{R}^d refers to a nonempty, connected open subset of the Euclidean space \mathbb{R}^d . We introduce the following notation for spaces of real-valued functions on a domain $D \subset \mathbb{R}^d$. The space $\mathcal{L}^p(D)$, $p \in [1, \infty)$, stands for the usual Banach space of (equivalence classes) of measurable functions f satisfying $\int_D |f(x)|^p dx < \infty$, and $\mathcal{L}^\infty(D)$ is the Banach space of functions that are essentially bounded in D . The space $\mathcal{C}^k(D)$ ($\mathcal{C}^\infty(D)$) refers to the class of all functions whose partial derivatives up to order k (of any order) exist and are continuous. The standard Sobolev space of functions on D whose generalized derivatives up to order k are in $\mathcal{L}^p(D)$, equipped with its natural norm, is denoted by $\mathcal{W}^{k,p}(D)$, $k \geq 0$, $p \geq 1$.

We adopt the notation $\partial_t \triangleq \frac{\partial}{\partial t}$, and for $i, j \in \mathbb{N}$, $\partial_i \triangleq \frac{\partial}{\partial x_i}$ and $\partial_{ij} \triangleq \frac{\partial^2}{\partial x_i \partial x_j}$. We often use the standard summation rule that repeated subscripts and superscripts are summed from 1 through d .

3.4. Some Facts from Parabolic Equations. For a nonnegative multi-index $\alpha = (\alpha_1, \dots, \alpha_d)$ we let $D^\alpha \triangleq \partial_1^{\alpha_1} \dots \partial_d^{\alpha_d}$. Let Q be a domain in $\mathbb{R}_+ \times \mathbb{R}^d$. Recall that $\mathcal{C}^{r,k+2r}(Q)$ stands for the set of bounded continuous functions $\varphi(t, x)$ defined on Q such that the derivatives $D^\alpha \partial_t^\ell \varphi$ are bounded and continuous in Q for

$$|\alpha| + 2\ell \leq k + 2r, \quad \ell \leq r. \quad (3.4)$$

For $\varphi \in \mathcal{C}^{r,k+2r}(\bar{Q})$ and $p \in [1, \infty)$, define

$$\|\varphi\|_{\mathcal{W}^{r,k+2r,p}(Q)} \triangleq \sum_{\substack{|\alpha| \leq k+2(r-\ell) \\ \ell \leq r}} \|D^\alpha \partial_t^\ell \varphi\|_{\mathcal{L}^p(Q)}.$$

The *parabolic Sobolev space* $\mathcal{W}^{r,k+2r,p}(Q)$ is the subspace of $\mathcal{L}^p(Q)$ which consists of those functions φ for which there exists a sequence φ_n in $\mathcal{C}^{r,k+2r}(\bar{Q})$ such that $\|\varphi_n - \varphi\|_{\mathcal{L}^p(Q)} \rightarrow 0$ as $n \rightarrow \infty$ and

$$\|D^\alpha \partial_t^\ell \varphi_n - D^\alpha \partial_t^\ell \varphi_m\|_{\mathcal{L}^p(Q)} \xrightarrow{n, m \rightarrow \infty} 0,$$

for all α and ℓ satisfying (3.4). In this way the Sobolev derivatives $D^\alpha \partial_t^\ell \varphi$ are well defined as functions in $\mathcal{L}^p(Q)$ and $\mathcal{W}^{r,k+2r,p}(Q)$ is a Banach space under the norm introduced.

Let $r : \mathbb{R}^d \times \mathbb{U}$ be a nonnegative continuous function which is locally Lipschitz continuous in x uniformly in $u \in \mathbb{U}$. Let κ_R be a Lipschitz constant of r over B_R .

We next review some standard estimates for solutions of equations of the form

$$-\partial_t \varphi(t, x) + \min_{u \in \mathbb{U}} [L^u \varphi(t, x) + r(x, u)] = f(t, x) \quad (3.5)$$

and

$$-\partial_t \varphi(t, x) + L^{v_t} \varphi(t, x) = g(t, x). \quad (3.6)$$

Note that if v is a measurable selector from the minimizer in (3.5) then the quasilinear equation (3.5) transforms to the linear equation (3.6), which in fact takes the particular form

$$-\partial_t \varphi(t, x) + a^{ij} \partial_{ij} \varphi(t, x) + H(D\varphi, x) = f(t, x), \quad (3.7)$$

where H is Lipschitz continuous in its arguments.

For $R > 0$ and $0 \leq T' < T$ define $B_R^{T', T} \triangleq (T', T) \times B_R$. Let $g \in \mathcal{W}^{0, k, p}(B_R^{0, T})$ and suppose that $\varphi \in \mathcal{W}^{1, 2, p}(B_R^{0, T})$ is a solution of (3.6). Then for any $R' \in (0, R)$ and $T' \in (0, T)$ it holds that $\varphi \in \mathcal{W}^{1, 2+k, p}(B_{R'}^{T', T})$ and there exists a constant $C_1 = C_1(R', R, T', T, k, d, \kappa_R, \kappa_1, p)$ such that

$$\|\varphi\|_{\mathcal{W}^{1, 2+k, p}(B_{R'}^{T', T})} \leq C_1 \left(\|g\|_{\mathcal{W}^{0, k, p}(B_R^{0, T})} + \|\varphi\|_{\mathcal{L}^p(B_R^{0, T})} \right). \quad (3.8)$$

Combining (3.8) with the compactness of the imbedding of $\mathcal{W}^{2, p}(B_R) \hookrightarrow \mathcal{C}^1(\bar{B}_R)$, for $p > d$, and the interpolation inequality, we conclude by using (3.7) that if $f \in \mathcal{W}^{0, 1, p}(B_R^{0, T})$, then $\varphi \in \mathcal{C}^{1, 2}(B_{R'}^{T', T})$ and

$$\max_{|\alpha| \leq 2} \sup_{B_{R'}^{T', T}} |D^\alpha \varphi| \leq C_2 \left(\|f\|_{\mathcal{W}^{0, k, p}(B_R^{0, T})} + \|\varphi\|_{\mathcal{L}^p(B_R^{0, T})} \right), \quad (3.9)$$

where C_2 depends on the parameters in C_1 . Moreover, if the derivatives $\partial_i f$ are bounded on $B_R^{0, T}$ then

$$\max_{|\alpha| \leq 1} \sup_{B_{R'}^{T', T}} |D^\alpha \partial_t \varphi| \leq C_3 \left(\max_{|\alpha| \leq 1} \sup_{B_{R'}^{T', T}} |D^\alpha f| + \|f\|_{\mathcal{W}^{0, k, p}(B_R^{0, T})} + \|\varphi\|_{\mathcal{L}^p(B_R^{0, T})} \right). \quad (3.10)$$

These estimates can be found in [Kry08, Chapter 5].

4. Main Results.

4.1. Proof of Theorem 2.3. The proof of Theorem 2.3 involves several intermediate results. For a subset Q of $\mathbb{R}_+ \times \mathbb{R}^d$, by abuse of notation, we let $\mathcal{C}_V(Q)$ denote the Banach space of functions in $\mathcal{C}(Q)$ with norm

$$\|f\|_V \triangleq \sup_{(t, x) \in Q} \left| \frac{f(t, x)}{V(x)} \right|.$$

Let $\mathbb{R}_T^d \triangleq [0, T] \times \mathbb{R}^d$. We next show that (2.13) has a unique solution in $\mathcal{C}_V(\mathbb{R}_T^d) \cap \mathcal{C}^{1, 2}(\mathbb{R}_T^d)$, for any $T > 0$.

LEMMA 4.1. *For each $V_0 \in \mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$, there exists a unique solution $V \in \mathcal{C}_V(\mathbb{R}_T^d) \cap \mathcal{C}^{1, 2}(\mathbb{R}_T^d)$, for any $T > 0$.*

Proof. We first show that if $g : [0, T] \rightarrow \mathbb{R}^d$ is a bounded continuous function, then

$$\partial_t \varphi(t, x) = \min_{u \in \mathbb{U}} [L^u \varphi(t, x) + r(x, u)] - g(t), \quad \varphi(0, x) = V_0(x) \quad (4.1)$$

has a unique solution in $\mathcal{C}_{\mathcal{V}}(\mathbb{R}_T^d) \cap \mathcal{C}^{1,2}(\mathbb{R}_T^d)$.

Let r^n denote the truncation of r , i.e., $r^n(c, u) \triangleq n \wedge r(x, u)$. Let τ_R denote the first exit time from the ball of radius R centered at the origin in \mathbb{R}^d , and let $\psi_R : \mathbb{R}^d \rightarrow [0, 1]$ be a smooth function which satisfies $\psi_R(x) = 1$ for $|x| \leq R/2$ and $\psi_R(x) = 0$ for $|x| \geq 3R/4$. Then the boundary value problem

$$\begin{aligned} \partial_t \varphi_{n,R}(t, x) &= \min_{u \in \mathbb{U}} [L^u \varphi_{n,R}(t, x) + r^n(x, u)] - g(t), \\ \varphi_{n,R}(0, x) &= V_0(x) \psi_R(x), \quad \varphi_{n,R}(t, \cdot)|_{\partial B_R} = 0 \quad \forall t \geq 0, \end{aligned} \quad (4.2)$$

has a unique solution in $\mathcal{C}^{1,2}(\mathbb{R}_T^d)$. This solution has the stochastic representation

$$\varphi_{n,R}(t, x) = \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[V_0(X_t) \psi_R(X_t) \mathbb{I}\{t < \tau_R\} + \int_0^{t \wedge \tau_R} (r^n(X_s, U_s) - g(s)) \, ds \right], \quad (4.3)$$

where \mathbb{I} denotes the indicator function. Since

$$\begin{aligned} \int_0^t \mathbb{E}_x^U [r^n(X_s, U_s)] \, ds &\leq c_2 \int_0^t \mathbb{E}_x^U [\mathcal{V}(X_s)] \, ds \\ &\leq \frac{c_2}{c_1} (c_0 t + \mathcal{V}(x)) \quad \forall U \in \mathfrak{U}, \end{aligned}$$

and $V_0 \in \mathcal{C}_{\mathcal{V}}(\mathbb{R}^d)$, we obtain

$$\varphi_{n,R}(t, x) \leq c_3(1 + \mathcal{V}(x)) + \frac{c_2}{c_1} (c_0 t + \mathcal{V}(x)) + \|g\|_{\mathcal{L}^1([0, T])} \quad (4.4)$$

for some constant $c_3 > 0$. Also by (4.3) we have

$$\varphi_{n,R}(t, x) \geq -c_3(1 + \mathcal{V}(x)) - \int_0^t g(s) \, ds, \quad (4.5)$$

and it follows that for any fixed g and V_0 , the solution $\varphi_{n,R}$ is bounded in $\mathcal{C}_{\mathcal{V}}(\mathbb{R}_T^d)$ uniformly in $R > 0$ and $n \in \mathbb{N}$. The interior estimates of solutions of (4.2) (see [LSU67, p. 342 and p. 351]) allow us to take limits as $R \rightarrow \infty$ (along some subsequence) to obtain a solution $\varphi_n \in \mathcal{C}^{1,2}(\mathbb{R}_T^d)$ to

$$\partial_t \varphi_n(t, x) = \min_{u \in \mathbb{U}} [L^u \varphi_n(t, x) + r^n(x, u)] - g(t), \quad \varphi_n(0, x) = V_0(x), \quad (4.6)$$

which naturally satisfies the bounds in (4.4)–(4.5). Using again the interior estimates of solutions to (4.6) we can let $n \rightarrow \infty$ to obtain in the limit a solution $\varphi \in \mathcal{C}_{\mathcal{V}}(\mathbb{R}_T^d) \cap \mathcal{C}^{1,2}(\mathbb{R}_T^d)$ to (4.1). Showing uniqueness of this solution is standard. Let φ and φ' be such solutions of (4.1) corresponding to g

and g' , respectively. Using the inequality $|\inf A - \inf B| \leq \sup |A - B|$ we have

$$\begin{aligned} \sup_{t \in [0, T]} |\varphi(t, 0) - \varphi'(t, 0)| &\leq \sup_{t \in [0, T]} \left| \inf_{U \in \mathfrak{U}} \mathbb{E}_0^U \left[V_0(X_t) + \int_0^t (r(X_s, U_s) - g(s)) \, ds \right] \right. \\ &\quad \left. - \inf_{U \in \mathfrak{U}} \mathbb{E}_0^U \left[V_0(X_t) + \int_0^t (r(X_s, U_s) - g'(s)) \, ds \right] \right| \\ &\leq \sup_{t \in [0, T]} \left| \int_0^t [g(s) - g'(s)] \, ds \right| \\ &\leq T \sup_{t \in [0, T]} |g(t) - g'(t)|. \end{aligned}$$

Hence for $T < 1$ the map $g(\cdot) \mapsto \varphi(\cdot, 0)$ is a contraction thus asserting the existence of a solution to (2.13) in $\mathcal{C}_{\mathcal{V}}(\mathbb{R}_T^d) \cap \mathcal{C}^{1,2}(\mathbb{R}_T^d)$, for $T < 1$. Concatenating intervals $[0, T]$, $[T, 2T]$, \dots , with $T < 1$, we obtain such a solution of (2.13) for any $T > 0$. Uniqueness is again standard. \square

The next two lemmas concern estimates for the solutions of the (RVI) and the (VI).

LEMMA 4.2. *For each $V_0 \in \mathcal{C}_{\mathcal{V}}(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$, the solution \bar{V} of (2.14) satisfies the bound*

$$|V^*(x) - \bar{V}(t, x)| \leq \|V^* - V_0\|_{\mathcal{V}} \left(\frac{c_0}{c_1} + \mathcal{V}(x)e^{-c_1 t} \right) \quad \forall x \in \mathbb{R}^d, \quad \forall t \geq 0. \quad (4.7)$$

Proof. Let v^* be a measurable selector from the minimizer in (2.11). Then

$$-\partial_t(V^* - \bar{V}) + L^{v^*}(V^* - \bar{V}) \leq 0 \quad (4.8)$$

from which, by an application of Itô's formula to $V^*(X_s) - \bar{V}(t - s, X_s)$, $s \in [0, t]$, it follows that

$$\mathbb{E}_x^{v^*} [V^*(X_t) - V_0(X_t)] \leq V^*(x) - \bar{V}(t, x). \quad (4.9)$$

On the other hand, if \bar{v} is a measurable selector from the minimizer in (2.14), then

$$-\partial_t(V^* - \bar{V}) + L^{\bar{v}}(V^* - \bar{V}) \geq 0,$$

and we obtain

$$V^*(x) - \bar{V}(t, x) \leq \mathbb{E}_x^{\bar{v}} [V^*(X_t) - V_0(X_t)]. \quad (4.10)$$

Since V^* and V_0 are in $\mathcal{C}_{\mathcal{V}}(\mathbb{R}^d)$, (4.7) follows by (2.10) and (4.9)–(4.10). \square

Remark 4.3. *Note that the Markov control associated with a measurable selector \bar{v} from the minimizer in (2.14) is computed ‘backward’ in time. Hence the control applied to the process X considered in (4.10) is the Markov control $U(s, x) = \bar{v}(t - s, x)$, $0 \leq s \leq t$, where \bar{v} solves*

$$\partial_t \bar{V}(t, x) = a^{ij}(x) \partial_{ij} \bar{V}(t, x) + b^i(x, \bar{v}(t, x)) \partial_i \bar{V}(t, x) + r(x, \bar{v}(t, x)) - \beta.$$

LEMMA 4.4. *If $\bar{V}(0, x) = V(0, x) = V_0(x)$ for some $V_0 \in \mathcal{C}_{\mathcal{V}}(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$, then the solutions V and \bar{V} of (2.13) and (2.14), respectively, satisfy*

$$V(t, x) - V(t, 0) = \bar{V}(t, x) - \bar{V}(t, 0) \quad (4.11)$$

$$V(t, x) = \bar{V}(t, x) - e^{-t} \int_0^t e^s \bar{V}(s, 0) \, ds + \beta(1 - e^{-t}) \quad (4.12)$$

for all $x \in \mathbb{R}^d$ and all $t \geq 0$.

Proof. By (2.13) and (2.14) we have

$$V(t, x) = \inf_{U \in \mathfrak{U}} \left(\int_0^t \mathbb{E}_x^U [r(X_s, U_s)] \, ds + \mathbb{E}_x^U [V_0(X_t)] \right) - \int_0^t V(s, 0) \, ds \quad (4.13a)$$

$$\bar{V}(t, x) = \inf_{U \in \mathfrak{U}} \left(\int_0^t \mathbb{E}_x^U [r(X_s, U_s)] \, ds + \mathbb{E}_x^U [V_0(X_t)] \right) - \beta t. \quad (4.13b)$$

Hence (4.11) follows by (4.13a)–(4.13b). Again by (4.13a)–(4.13b) we have

$$V(t, 0) - \beta + \int_0^t (V(s, 0) - \beta) \, ds = \bar{V}(t, 0) - \beta, \quad (4.14)$$

and solving (4.14) we obtain

$$V(t, 0) = \bar{V}(t, 0) - e^{-t} \int_0^t e^s \bar{V}(s, 0) \, ds + \beta(1 - e^{-t}),$$

which combined with (4.11) yields (4.12). \square

Next we show that the solution \bar{V} of the (VI) converges as $t \rightarrow \infty$ for any initial condition V_0 .

THEOREM 4.5. *For each $V_0 \in \mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$, $\bar{V}(t, x) \rightarrow V^*(x) + c$ as $t \rightarrow \infty$, for some $c \in \mathbb{R}$ which depends on V_0 .*

Proof. We view the solutions of (2.14) as a semiflow on $\mathcal{H} = \mathcal{Y} = \mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$, also letting $\mathcal{G} = \mathcal{C}_V(\mathbb{R}^d)$, and apply Theorem 3.1. We equip $\mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$ with a complete metric, for example by letting

$$d(f, g) \triangleq \|f - g\|_{\mathcal{C}_V(\mathbb{R}^d)} + \sum_{n=1}^{\infty} \frac{1}{2^n} \max \left(1, \|f - g\|_{\mathcal{C}^2(B_n)} \right),$$

where B_n denotes the ball of radius n centered at the origin in \mathbb{R}^d and

$$\|f\|_{\mathcal{C}^2(B)} \triangleq \sum_{|\alpha| \leq 2} \sup_B |D^\alpha f|.$$

Hypothesis (i) of Theorem 3.1 is clearly satisfied. Let $\Phi_t(V_0) : \mathbb{R}^d \rightarrow \mathbb{R}$ denote the solution of (2.14) corresponding to $V_0 \in \mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$. Let $\mathcal{E} \triangleq \{V^* + c : c \in \mathbb{R}\} \subset \mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$, i.e., the set of equilibria of this semiflow. Note the following:

- (a) for each $V_0 \in \mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$, $\Phi_t(V_0)$ is bounded in $\mathcal{C}_V(\mathbb{R}^d)$ by (4.7). Also the second order partial derivatives of $\Phi_t(V_0)$ are locally equicontinuous in x , uniformly in $t \geq T$ for some $T > 0$ (this requires a slight improvement of (3.9), adding Hölder continuity which is standard [LSU67, Theorem 5.1]). Hence, every subsequence $\Phi_{t_n}(V_0)$ contains a further subsequence that converges in $\mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$, which, in turn, implies that the orbit $\{\Phi_t(V_0) : t \in \mathbb{R}_+\}$ has a compact closure in $\mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$.
- (b) If $\{V_0^n\} \subset \mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$ is a monotone sequence such that $V_0^n \rightarrow V_0 \in \mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$ as $n \rightarrow \infty$, then by (4.7) the set $\{\cup_{n \in \mathbb{N}} \Phi_t(V_0^n) : t \in \mathbb{R}_+\}$ is bounded in $\mathcal{C}_V(\mathbb{R}^d)$. Hence it has locally Hölder equicontinuous second order partial derivatives in x , which implies that it has a compact closure in $\mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$. In particular, the set $\{\cup_{n \in \mathbb{N}} \omega(V_0^n)\}$ has compact closure in $\mathcal{C}_V(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$.

Hence assumptions (ii) and (iii) of Theorem 3.1 are satisfied.

Consider the partial order relation \preceq on $\mathcal{C}_{\mathcal{V}}(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$ induced by the positive cone $\mathcal{C}_{\mathcal{V}}(\mathbb{R}^d)_+$ in $\mathcal{C}_{\mathcal{V}}(\mathbb{R}^d)$. Indeed, if $V_0 \prec V'_0$ then (2.14) yields

$$\mathbb{E}_x^{v'} [V'_0(X_t) - V_0(X_t)] \leq \Phi_t(V'_0)(x) - \Phi_t(V_0)(x) \quad \forall t > 0, \quad \forall x \in \mathbb{R}^d, \quad (4.15)$$

where v' is a Markov control associated with a measurable selector from the minimizer in (2.14) corresponding to the solution starting at V'_0 (see Remark 4.3). It follows from (4.15) and the fact that the support of the transition probabilities of the controlled process is the entire space \mathbb{R}^d that if $V_0 \prec V'_0$, then $\Phi_t(V_0) \prec \Phi_t(V'_0)$ for all $t > 0$, or in other words that the semiflow Φ is strongly monotone on $\mathcal{C}_{\mathcal{V}}(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$. As mentioned in Section 3.1 the semiflow is then strongly order-preserving. Since \mathcal{E} is totally ordered it follows by Theorem 3.1 that $\mathcal{H} = \overline{\text{int}(\mathcal{E})}$.

It remains to show that \mathcal{E} is closed. Note that

$$\begin{aligned} |\Phi_t(V_0)(x) - \Phi_t(V'_0)(x)| &\leq \sup_{U \in \mathfrak{U}} |\mathbb{E}_x^U [V_0(X_t) - V'_0(X_t)]| \\ &\leq \|V_0 - V'_0\|_{\mathcal{V}} \left(\sup_{U \in \mathfrak{U}} \mathbb{E}_x^U [\mathcal{V}(X_t)] \right), \quad t \geq 0, \end{aligned}$$

Hence by (2.10) we have

$$\begin{aligned} \|\Phi_t(V_0) - \Phi_t(V'_0)\|_{\mathcal{V}} &\leq \left(\sup_{x \in \mathbb{R}^d} \sup_{U \in \mathfrak{U}} \frac{\mathbb{E}_x^U [\mathcal{V}(X_t)]}{\mathcal{V}(x)} \right) \|V_0 - V'_0\|_{\mathcal{V}} \\ &\leq \left(1 + \frac{c_0}{c_1} \right) \|V_0 - V'_0\|_{\mathcal{V}}, \quad t \geq 0. \end{aligned}$$

This shows in particular that if V_0^n is a Cauchy sequence of convergent points in $\mathcal{C}_{\mathcal{V}}(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$, then $f_n \triangleq \omega(V_0^n)$ converges in $\mathcal{C}_{\mathcal{V}}(\mathbb{R}^d)$ as $n \rightarrow \infty$. Suppose that $V_0 \in \mathcal{E}^c$. Since \mathcal{E} is dense, there exists $\{V_0^n\} \subset \mathcal{E}$ such that $V_0^n \rightarrow V_0$ as $n \rightarrow \infty$. Let $f \triangleq \lim_{n \rightarrow \infty} \omega(V_0^n)$. Since $V_0 \in \mathcal{E}^c$, then $\limsup_{t \rightarrow \infty} d(\Phi_t(V_0), f) > 0$. Moreover, since for some $T > 0$ the set $\{\Phi_t(V_0) : t > T\}$ is precompact in $\mathcal{C}_{\mathcal{V}}(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$ there exists $f' \in \mathcal{C}_{\mathcal{V}}(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$, $f \neq f'$ and a sequence t'_n such that $\Phi_{t'_n}(V_0) \rightarrow f'$ as $n \rightarrow \infty$. On the other hand, we can find a sequence t_n such that $\sup_{t > t_n} \|\Phi_t(V_0^n) - f\|_{\mathcal{V}} \rightarrow 0$ in $\mathcal{C}_{\mathcal{V}}(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$ as $n \rightarrow \infty$. Therefore, for some subsequence $n(k) \uparrow \infty$, we have $\|\Phi_{t_{n(k)}}(V_0^k) - f\|_{\mathcal{V}} \rightarrow 0$ as $k \rightarrow \infty$. Therefore,

$$\begin{aligned} 0 &< \|f' - f\|_{\mathcal{V}} \\ &= \lim_{k \rightarrow \infty} \|\Phi_{t_{n(k)}}(V_0) - \Phi_{t_{n(k)}}(V_0^k)\|_{\mathcal{V}} \\ &\leq \left(1 + \frac{c_0}{c_1} \right) \lim_{k \rightarrow \infty} \|V_0 - V_0^k\|_{\mathcal{V}} \\ &= 0, \end{aligned}$$

yielding a contradiction. Thus we have shown that all points of $\mathcal{C}_{\mathcal{V}}(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$ are convergent and the proof is complete. \square

We are now ready for the proof of the main result.

Proof of Theorem 2.3. If we define $g(t) \triangleq V(t, x) - \bar{V}(t, x)$, then by (4.12) we have

$$g(t) = \int_0^t e^{s-t} (\beta - \bar{V}(s, 0)) \, ds$$

Since $\bar{V}(t, x) \rightarrow V^*(x) + c$ as $t \rightarrow \infty$ for each $x \in \mathbb{R}^d$ by Theorem 4.5, it follows that $g(t)$ converges to $\beta - V^*(0) - c = \beta - c$ as $t \rightarrow \infty$. Hence

$$\lim_{t \rightarrow \infty} V(t, x) = \lim_{t \rightarrow \infty} [g(t) + \bar{V}(t, x)] = V^*(x) + \beta \quad \forall x \in \mathbb{R}^d,$$

and the proof is complete. \square

4.2. An alternate proof of Theorem 4.5. Recall that v^* is an optimal stationary Markov control. Let μ_{v^*} be the corresponding invariant probability distribution, and let X_t^* , $t \in \mathbb{R}$, be a stationary solution of (2.1) under the control v^* such that the law of X_t^* is μ_{v^*} for all $t \in \mathbb{R}$. Let $\mathfrak{F}_t \triangleq \sigma(X_s : -\infty < s \leq t)$ and

$$\Psi(t, x) \triangleq \bar{V}(t, x) - V^*(x).$$

By (4.8) we have

$$-\partial_t \Psi(t, x) + L^{v^*} \Psi(t, x) \geq 0.$$

Therefore the process

$$M_t \triangleq \Psi(t, X_{-t}^*), \quad t \in [0, \infty),$$

is a reverse (\mathfrak{F}_{-t}) -supermartingale. Also, by (4.7) there exists a constant C_0 such that $\mathbb{E}[|M_t|] \leq C_0$, for all $t \in [0, \infty)$. We argue by contradiction. Suppose that $\bar{V}(t, \cdot) - V^*(\cdot)$ does not converge to a constant as $t \rightarrow \infty$. Then, there must exist constants $a < b$, a ball $D \subset \mathbb{R}^d$ and a pair of sequences $\{t_n\} \subset \mathbb{R}_+$ and $\{x_n\} \subset D$, $n \in \mathbb{N}$, such that

$$\Psi(t_{2k-1}, x_{2k-1}) \leq a, \quad \Psi(t_{2k}, x_{2k}) \geq b \quad \forall k \in \mathbb{N}. \quad (4.16)$$

Let $B_r(x)$ denote the open ball of radius r centered at $x \in \mathbb{R}^d$. Since $\bar{V}(\cdot, t) - V^*(\cdot)$ is uniformly equicontinuous on any bounded domain, there exists $r > 0$, such that if $x \in D$, then

$$|\Psi(t, x) - \Psi(t, y)| \leq \frac{b-a}{4} \quad \forall y \in B_{2r}(x). \quad (4.17)$$

Let $\mathcal{G} = \{G_i : 1 \leq i \leq N\}$ be a finite open cover of D with balls of radius r . Since \mathcal{G} is finite, an infinite number of terms of the sequences $\{x_{2k-1} : k \in \mathbb{N}\}$ and $\{x_{2k} : k \in \mathbb{N}\}$ lie in some elements G' and G'' of \mathcal{G} , respectively. Dropping to a subsequence of $\{x_k\}$, which is also denoted as $\{x_k\}$, it follows by (4.16)–(4.17) that

$$\begin{aligned} \Psi(t_{2k-1}, x) &\leq a' \triangleq \frac{3a+b}{4} & \forall x \in G' \\ \Psi(t_{2k}, x) &\geq b' \triangleq \frac{a+3b}{4} & \forall x \in G'' \end{aligned} \quad (4.18)$$

for all $k = 1, 2, \dots$. Without loss of generality we can also assume that the time sequence t_n satisfies $t_{n+1} - t_n > \gamma_0 > 0$. The convergence of the transition probability under the control v^* implies that for some constant $\varepsilon_0 > 0$

$$\mathbb{P}_x^{v^*}(X_t^* \in B_r(y)) \geq \varepsilon_0 \quad \forall x, y \in D, \quad \forall t > \gamma_0. \quad (4.19)$$

It follows by (4.18)–(4.19) that

$$\mathbb{P}_{\mu_{v^*}}^{v^*} \left(\sum_{k \in \mathbb{N}} \mathbb{I}\{X_{t_{2k-1}}^* \in G'\} < \infty \right) = 0 \quad \text{and} \quad \mathbb{P}_{\mu_{v^*}}^{v^*} \left(\sum_{k \in \mathbb{N}} \mathbb{I}\{X_{t_{2k}}^* \in G''\} < \infty \right) = 0$$

Therefore,

$$\mathbb{P}_{\mu_{v^*}}^{v^*} \left(\sum_{k \in \mathbb{N}} \mathbb{I}\{X_{t_{2k-1}}^* \in G'\} = \infty, \sum_{k \in \mathbb{N}} \mathbb{I}\{X_{t_{2k}}^* \in G''\} = \infty \right) = 1. \quad (4.20)$$

Therefore if ν is the number of upcrossings of $[a', b']$ by M then (4.18) and (4.20) imply that $\mathbb{P}_{\mu_{v^*}}^{v^*}(\nu < \infty) = 0$. However by the reverse submartingale upcrossings inequality $\mathbb{E}_x^{v^*}[\nu] < \infty$ which gives a contradiction and the proof is complete.

5. Conclusions. We have proposed a nonlinear parabolic PDE that serves as a continuous time, continuous state space analog of the relative value iteration scheme for solving the ergodic dynamic programming equation in finite state problems. This was done under a uniform stability condition in terms of an associated Lyapunov function.

These results suggest several future directions:

1. An important class of ergodic control problems is one wherein instability is possible, but is heavily penalized by using a ‘near-monotone’ (see [ABG11, Chapter 3] for a definition) running cost. It would be both interesting and important to extend the above results to this case as it covers several important applications.
2. While the foregoing seems to extend easily to two-person zero-sum stochastic differential games with ergodic payoffs, it would be of great interest to do the same for interesting classes of non-cooperative games with ergodic payoffs.
3. Rate of convergence results, computational aspects, and convergence under subgeometric ergodicity are also open issues.

REFERENCES

- [ABB01] J. Abounadi, D. P. Bertsekas, and V. S. Borkar, *Learning algorithms for Markov decision processes with average cost*, SIAM J. Control Optim. **40** (2001), no. 3, 681–698.
- [ABG11] A. Arapostathis, V. S. Borkar, and M. K. Ghosh, *Ergodic control of diffusion processes*, Encyclopedia of Mathematics and its Applications, vol. 143, Cambridge University Press, Cambridge, 2011.
- [AF99] Y. Aviv and A. Federgruen, *The value iteration method for countable state Markov decision processes*, Oper. Res. Lett. **24** (1999), no. 5, 223–234.
- [Ber98] D. P. Bertsekas, *A new value iteration method for the average cost dynamic programming problem*, SIAM J. Control Optim. **36** (1998), no. 2, 742–759.
- [BKR01] V. I. Bogachev, N. V. Krylov, and M. Röckner, *On regularity of transition probabilities and invariant measures of singular diffusions under minimal conditions*, Comm. Partial Differential Equations **26** (2001), no. 11–12, 2037–2080.
- [DMT95] D. Down, S. P. Meyn, and R. L. Tweedie, *Exponential and uniform ergodicity of Markov processes*, Ann. Probab. **23** (1995), no. 4, 1671–1691.
- [DY79] E. B. Dynkin and A. A. Yushkevich, *Controlled Markov processes*, Grundlehren der mathematischen Wissenschaften, vol. 235, Springer-Verlag, New York, 1979.
- [FR05] G. Fort and G. O. Roberts, *Subgeometric ergodicity of strong Markov processes*, Ann. Appl. Probab. **15** (2005), no. 2, 1565–1589.
- [GK96] I. Gyöngy and N. Krylov, *Existence of strong solutions for Itô's stochastic equations via approximations*, Probab. Theory Related Fields **105** (1996), no. 2, 143–158.
- [GS72] I. I. Gihman and A. V. Skorohod, *Stochastic differential equations*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 72, Springer-Verlag, Berlin, 1972.
- [Has60] R. Z. Has'minskiĭ, *Ergodic properties of recurrent diffusion processes and stabilization of the solution of the Cauchy problem for parabolic equations*, Theory Probab. Appl. **5** (1960), no. 2, 179–196.
- [JFHL09] H. Jasso-Fuentes and O. Hernández-Lerma, *Ergodic control, bias, and sensitive discount optimality for Markov diffusion processes*, Stoch. Anal. Appl. **27** (2009), no. 2, 363–385.
- [Kry80] N. V. Krylov, *Controlled diffusion processes*, Applications of Mathematics, vol. 14, Springer-Verlag, New York, 1980.
- [Kry08] ———, *Lectures on elliptic and parabolic equations in Sobolev spaces*, Graduate Studies in Mathematics, vol. 96, American Mathematical Society, Providence, RI, 2008.
- [LSU67] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural'ceva, *Linear and quasilinear equations of parabolic type*, Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23, American Mathematical Society, Providence, R.I., 1967.
- [Put94] M. L. Puterman, *Markov decision processes: Discrete stochastic dynamic programming*, John Wiley & Sons Inc., New York, 1994.
- [SLKJ10] O. Shlakhter, C.-G. Lee, D. Khmelev, and N. Jaber, *Acceleration operators in the value iteration algorithms for Markov decision processes*, Oper. Res. **58** (2010), no. 1, 193–202.
- [Smi95] H. L. Smith, *Monotone dynamical systems*, Mathematical Surveys and Monographs, vol. 41, American Mathematical Society, Providence, RI, 1995.
- [Sta99] W. Stannat, *(Nonsymmetric) Dirichlet operators on L^1 : existence, uniqueness and associated Markov processes*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **28** (1999), no. 1, 99–140.
- [Whi63] D. J. White, *Dynamic programming, Markov chains, and the method of successive approximations*, J. Math. Anal. Appl. **6** (1963), 373–376.